UNCLASSIFIED

AFCCDD TN 61-33

RF Project 1080
Technical Report 60

TECHNICAL

REPORT

by

THE OHIO STATE UNIVERSITY
RESEARCH FOUNDATION

COLUMBUS 12, OHIO

TO: OPERATIONAL APPLICATIONS OFFICE
COMMAND AND CONTROL DEVELOPMENT DIVISION
AIR RESEARCH AND DEVELOPMENT COMMAND
UNITED STATES AIR FORCE
LAURENCE G. HANSCOM FIELD
BEDFORD, MASSACHUSETTS
Contract No. AF 19(604)-6179

ON: INTELLIGIBILITY AND CONFUSABILITY OF VOWELS AND DIPHTHONGS UNDER VARIOUS CONDITIONS OF QUIET AND NOISE

Submitted by: Henry M. Moser
John F. Michel
Wallace C. Fotheringham

Date: March 1961

AFCCDD TN 61-33

RF Project 1080
Technical Report 60

TECHNICAL

REPORT

by

THE OHIO STATE UNIVERSITY
RESEARCH FOUNDATION

COLUMBUS 12, OHIO

TO: OPERATIONAL APPLICATIONS OFFICE
COMMAND AND CONTROL DEVELOPMENT DIVISION
AIR RESEARCH AND DEVELOPMENT COMMAND
UNITED STATES AIR FORCE
LAURENCE G. HANSCOM FIELD
BEDFORD, MASSACHUSETTS
Contract No. AF 19(604)-6179

ON: INTELLIGIBILITY AND CONFUSABILITY OF VOWELS AND DIPHTHONGS UNDER VARIOUS CONDITIONS OF QUIET AND NOISE

Submitted by: Henry M. Moser
John F. Michel
Wallace C. Fotheringham

Date: March 1961

TABLE OF CONTENTS

		PAGE
I	INTRODUCTION	1
II	REVIEW OF PREVIOUS RESEARCH	1
III	PURPOSES OF THE STUDY	2
IV	METHOD	2
	Stimuli and Recording Procedure	2
	Speakers	3
	Listeners	3
	Dependent and Independent Variables	3
	Test Administration	4
V	ANALYSIS OF DATA AND RESULTS	6
VI	CONCLUSIONS	17
	BIBLIOGRAPHY	18

ABSTRACT

Sixteen English vowels and diphthongs were recorded by four male and four female speakers. For each vowel and diphthong, 120 responses were obtained from a panel of six phonetically trained listeners at several levels of noise and quiet; each of five levels of noise were matched for listening difficulty with five levels of quiet. The levels of difficulty ranged from approximately 25 per cent to 85 per cent correct.

The results on vowel-diphthong intelligibility support the conclusions that vowels and diphthongs (1) are significantly different in intelligibility; (2) have a fairly stable order of intelligibility, similar in noise and quiet and among the eight speakers, though more stable among speakers of the same sex; (3) improve in intelligibility at different rates as listening conditions are improved; and (4) are more intelligible from male speakers.

In regard to vowel and diphthong confusability, results support the conclusions that (1) a great many significant confusions exist among the vowels and diphthongs under fairly difficult listening conditions, (2) at least one significant confusion exists for each vowel and diphthong, (3) each vowel and diphthong is a significant confusion for at least one other vowel or diphthong, and (4) confusions bear a reciprocal relationship to one another.

LIST OF TABLES

		PAGE
I	CORRECT RESPONSES AT DIFFERENT LEVELS OF QUIET (Q) AND NOISE (N)	7
II	NUMBER AND PERCENTAGE OF CORRECT IDENTIFICATIONS FOR 2400 PRESENTATIONS FOR EACH OF 16 VOWELS OR DIPHTHONGS	8
III	WILCOXIN T VALUES FOR THE DIFFERENCES IN INTELLIGIBILITY OF PAIRS OF VOWELS OR DIPHTHONGS	9
IV	CORRELATIONS OF VOWEL AND DIPHTHONG INTELLIGIBILITY VALUES BETWEEN FIVE LEVELS OF STIMULUS CLARITY	11
V	RESULTS OF WILCOXIN TEST OF THE DIFFERENCE IN INTELLIGIBILITY VALUES BETWEEN MALE AND FEMALE SPEAKERS UNDER VARIOUS CONDITIONS	12
VI	CONFUSION MATRIX FOR VOWELS IN QUIET (Q) AND NOISE (N)	13a
VII	RELATIONSHIPS BETWEEN RELATED CONFUSIONS FOR EACH OF 16 VOWELS OR DIPHTHONGS	14

FIGURES

1	PERCENTAGES OF CORRECT RESPONSES TO VOWEL STIMULI IN QUIET AND NOISE AT VARIOUS LEVELS OF LISTENING DIFFICULTY	5
2	PRINCIPAL CONFUSIONS IN QUIET	15
3	PRINCIPAL CONFUSIONS IN NOISE	16

I INTRODUCTION

Considerable research has been directed toward determining the intelligibility of the consonantal sounds of English. Very little has been done on the intelligibility of the English vowels, particularly at low intensities and relatively difficult signal-to-noise ratios.

Knowledge of the intelligibility of the vowels at such levels should permit more effective use of the English language in the selection or formation of words for use in communication. Speech via telephone or radio, and the development of an International Language for Aviation Communication, are particular areas in which accuate transmission under less than perfect listening conditions is needed.

Further, the study of vowels is important in that they are the loudest parts of words. At low intensities or when consonants are masked by noise,[1] the expectation is that the vowel or vowels might still be heard. Correct identification of the word might still occur. Consequently, the use of a word containing an intelligible vowel should be more satisfactory for communication than the use of the same consonantal structure with a less intelligible vowel. In this way, a possible "building block" approach might be employed in the construction of a vocabulary for use when optimum listening conditions are not assured.[2]

II REVIEW OF PREVIOUS RESEARCH

Difficulties in the auditory perception of an individual word often are related to faintness or indistinctness of the component speech sounds. Faintness and indistinctness, in turn, are associated with low intensities of the signal, the presence of masking noise, or high confusability between sounds. Regarding the vowels, studies by Sacia and Beck[3]; Black[4]; Fairbanks, House, and Stevens[5]; and Curry[6]; have been concerned with determining the range of relative intensities of the vowels in decibels. The results of these studies are not entirely consistent. For example, the study by Black correlates with the study by Sacia and Beck with a rank order correlation of .47, while the study by Fairbanks, House, and Stevens correlates with the Curry study with a rank order correlation of .87. The latter work was additionally concerned with the threshold identification of the vowels, and Curry found that vowels having greater intensity were not always the most easily identified. He concluded that intensity was not the only factor involved in vowel identification.

Siegenthaler,[7] in a study of sustained vowels at supra-threshold levels, reported that many of the sustained vowls tended to sound like [ʌ] whenever the initiations and conclusions of the vowels were removed. Koch[8] states that sustained vowels should not be used in communication because such steady sounds tend to lose their identification. Moser, Dreher, and O'Neill[9] investigated the masking of English monosyllabic words by prolonged vowel sounds. Prolonged vowel sounds were found to differ greatly in masking speech; vowels with concentration of energy in the 700-1000 cps range were the most effective masking agents; a monosyllable containing the same vowel as that employed as a masking agent was not more likely than others to be blotted out; monosyllables containing [ɔ] and [o] were most affected, those with [ʌ] the least; the rank order of vowel-masking effectiveness was [ɛ], [æ], [ɔ], [ʊ], [ɪ], [ɑ], [i], [o], and [u].

A number of studies have considered formant structure and position responsible for vowel identification. Tiffany[10] found a high correlation between formant position and threshold identification of vowels. Lehiste and Peterson[11], in a study of filtered vowels, found that only the first three formants were important, and that different vowels depended on different formants and the interaction of these formants for identification. Peterson collaborated with Barney[12] on methods of vowel study to determine the relative intelligibility of the vowels. Lists contained 10 monosyllabic words each beginning with [h] and ending with [d] and differing in the vowel. These were presented to listeners at a supra-liminal level of 70 db (re 0.0002 dyne/cm^2). They report: "Certain of the vowels ([i], [ɝ], [æ], and [u]) are generally better understood than others, possibly because they represent limit positions of the articulatory mechanism." They also found that, when observers confused one vowel with another, the two vowels nearly always had adjacent positions on the vowel loop; i.e., [ɪ] was taken for [ɛ] and [ɛ] was called either [ɪ] or [æ]. Edmonson and Horwitz[13] reported that vowel confusions are the result of the overlap of the first two formants, although for the most part vowels are recognized a high percentage of the time. Miller[14] used a 670-cycle low-pass filter to eliminate the second formant of 16 vowels and diphthongs presented in a consonantal [h - d] context. Confusions showed very closely the pattern that would be predicted from formant analysis. Since the filter would not affect temporal characteristics, Miller offers as the simplest explanation of the discrepancies the deduction that duration is an important feature. He concludes that there are at least three distinctive features for simple vowels: duration, frequency of the first formant, and frequency of the second formant. Pickett[15] studied the effects of various noise spectra on the intelligibility of vowels and found that all vowels are not affected in the same way by the same noise spectra. Significant shifts in vowel confusions occurred with changes in noise spectra, and these changes were consistent with the formant theory. The intelligibility of a sound under such conditions could be determined by its relative intensity or its characteristic phonetic nature.

III PURPOSES OF THIS STUDY

The purposes of this study were to determine: (1) the differences in intelligibility of isolated vowels and diphthongs at low intensities and relatively difficult signal-to-noise ratios; (2) the similarity of vowel and diphthong rank order intelligibility in quiet and noise; (3) the stability of vowel and diphthong rank order intelligibility as the intensity of the stimuli are successively increased; (4) the consistency of vowel and diphthong rank order intelligibility among various speakers; (5) the similarity of intelligibility values of the vowels and diphthongs between male and female speakers; (6) the principal confusions between the vowels and diphthongs under conditions of quiet and noise; and (7) the similarity of confusability values among related pairs of errors.

IV METHOD

Stimuli and Recording Procedure: Randomized lists of isolated vowels and diphthongs were recorded by eight speakers. Each speaker recorded a different randomization of 80 stimuli; each randomization contained five presentations of each vowel and diphthong. The 16 common vowels and diphthongs used were: [i], [ɪ], [e], [ɛ], [æ], [ɑ], [ɔ], [ʌ], [ɝ], [o], [ʊ], [u], [aɪ], [aʊ], [ɔɪ], and [ɪu].* Segments of a typical recording

*For those unfamiliar with phonetic symbols the above vowels and diphthongs are identified by underlining sounds in the following common words: HE, HIT, HAY, HECK, HAT, HOT, HAWK, HUT, HER, HOE, HOOK, WHO, HIGH, HOW, HOIST, and HUE.

follow:

"This is Speaker Number Five,
Number 1, write [ɑ],
Number 2, write [ɪ],
Number 3, write [aɪ],
....
Number 80, write [ɛ]."

In addition, an orientation and training tape was prepared; this provided an opportunity for listeners to become familiar with the voices of the eight speakers and the stimuli to be identified. It also provided the materials for a listener training program, and the means of determining the average auditory detection threshold of the listeners.

Recording was done with a tape recorder (Ampex 600), using a condenser microphone (Altec 21-B) positioned at the corner of the mouth, lightly touching the cheek. The original recordings were played through a laboratory signal-to-noise equalizer[16] and re-recorded. In this operation the word "write" of every carrier phrase was equated to within $\pm$ 1 db. However, the relationship of each stimulus to its introductory carrier phrase was maintained. For example, if the "write" was raised one or two db in intensity, the stimulus vowel was raised by the same amount. After all of the carrier phrases were equated, each was increased 10 db to ensure that the carrier words would be clear enough to prepare the listeners for the stimuli and enable them to write their responses in the appropriate spaces on the answer forms.

Speakers: Four males and four females of General American dialect were selected and trained as speakers. All were familiar with phonetics and were experienced as laboratory talkers. Detailed instructions were given to keep the carrier word and the stimulus as near to equal intensity as possible, and to make the stimuli equal in duration. Practice was given via speaking into a microphone connected to a tape recorder equipped with a VU (volume unit) meter. In this way, the speakers were able to monitor the level of the carrier word and to establish a more homogeneous pattern for use in the acutal recording, during which the VU meter was not used. Therefore, each speaker used his own natural feedback mechanism to maintain his vowels at their own natural level rather than at a constant energy level. It was assumed that, by keeping the vocal level constant, the inherent intelligibility of each vowel and diphthong would be maintained.

Listeners: Six research assistants, three male and three female, of the Psycholinguistics Laboratory at The Ohio State University served as listeners. All were trained listeners, had normal hearing, and had had at least one course in phonetics.

Dependent and Independent Variables: In accordance with the purposes of this study, two dependent variables (or properties) of vowels were of interest. These were intelligibility and confusability. Intelligibility was operationally defined as the percentage of correct identifications in the total number of choices made. Confusability was defined as the percentage of the total incorrect identifications in which a particular vowel or diphthong is substituted for the stimulus actually spoken; omissions were excluded from the computations.

The purposes of the study additionally called for the systematic variation of several independent variables. These involved differences in stimuli presented (16 vowels and diphthongs), differences in the intensity of stimulus presentation and differences in signal-to-noise ratios, differences in the sex of speakers, and differences in individuals as speakers.

The intensity of the stimulus presentation in quiet was varied to produce five levels. The first level of presentation was 4 db above the average detection threshold of the listeners. At this level a number of the stimuli were below threshold. In the original recording the vowels and diphthongs were produced at a constant vocal level and recorded on this basis rather than adjusted to a constant peak energy level; therefore, some of the stimuli may have had more energy than others even to the point of being supra-threshold at the lowest level of presentation. The second level of presentation was 8 db above the average detection threshold, the third level, 12 db, the fourth level, 16 db, and the fifth level, 20 db.

A careful attempt was made to equate the conditions in noise with the conditions in the quiet with respect to listening comprehension. The rationale employed was that if the number of correct answers to the training tape at level one in noise was approximately the same as at level one in quiet, this approximation would continue for all pairs of levels. A level in noise was selected on the basis of the previous experience of the laboratory staff to yield the same number of correct answers as were given at the most difficult level in quiet. This chosen level was a -10 db signal-to-noise ratio in reference to the stimuli. Succeeding signal-to-noise ratios were -6, -2, +2, and +6 db in that order. When the data were tabulated (see Table I), it was apparent that the percentages of correct responses in quiet and noise were quite satisfactorily equalized at each of the five levels of listening difficulty. This equating of difficulty can also be observed in Figure 1.

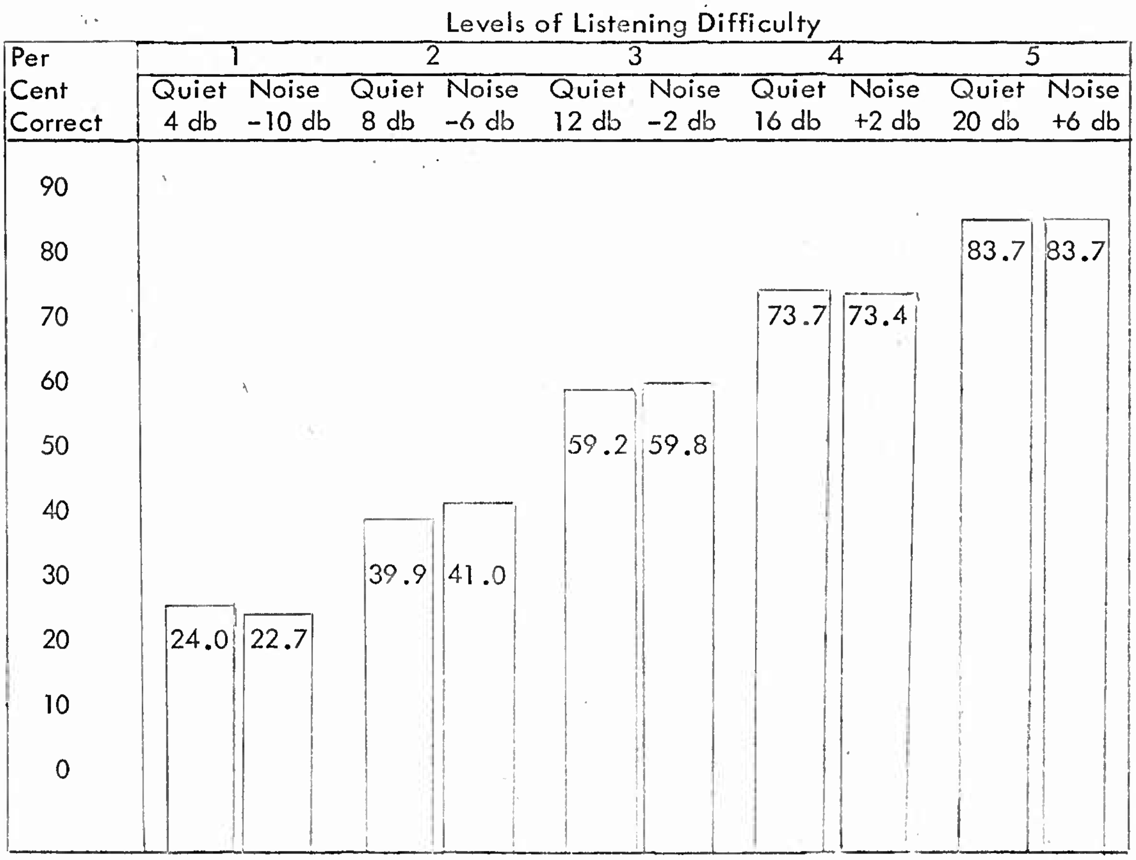
The signal-to-noise ratios were obtained in the following manner: the output of a flat noise generator (Grason-Stadler, Model 455-B) was fed into the laboratory signal-to-noise equalizer[16] to cause a certain deflection of the needle on the microammeter. The noise was filtered in the laboratory equalizer by a low-pass filter cutting off at 4500 cps with a slope of -18 db per octave. The speech signals were also fed into the equalizer where they were adjusted by means of vertical row of five indicator lights. When the intensity of the signal was sufficient to cause the lower four lights to glow, the peak voltage of the signal was very closely approximated to the rms voltage of the noise. Since all the carrier words had been equated earlier, the speech and noise voltages were mixed through attenuators set for the desired signal-to-noise ratio. After the signal-to-noise ratio was set relative to the carrier words, the carrier-to-noise ratio was adjusted so as to be 10 db greater than the vowel-to-noise ratio. In this way, the carrier words would be clear enough to enable the listeners to locate the appropriate spaces on the answer forms. In the testing in noise, the noise was maintained at the same level and the signal was adjusted in intensity to produce the various signal-to-noise ratios.

Test Administration: Prior to the actual testing, the listeners were acquainted with the 16 different stimuli and the speaker voices they would hear, and practice sessions were conducted to produce stable scores with the training tape. This same tape was also used to determine the average auditory detection threshold of the listeners to the stimuli, in both ascending and descending manner. The range of the mean detection thresholds for the individual listeners was 2 db.

Listeners were seated in a prefabricated sound-treated room (Industrial Acoustics Company, Model 403). The ambient noise level in the test chamber during testing was 30 db as measured with a sound-level meter (H.H. Scott, Model 410-A, C scale). Listening was monaural with the stimuli delivered to the preferred ear of the listener. Listeners transcribed their responses on specially prepared answer forms.

FIGURE 1

PERCENTAGES OF CORRECT RESPONSES TO VOWEL STIMULI IN QUIET AND NOISE AT VARIOUS LEVELS OF LISTENING DIFFICULTY



Tests in the quiet were administered through a tape recorder (Ampex 600) to six sets of headphones (Telephonics TDH-39, mounted in MX-41/AR cushions). The output of the recorder was fed to an attenuator (Hewlett-Packard), then into a mixing transformer (UTC, Model CG-137), and finally into a listening circuit containing the six headphones. The accuracy of the attenuation response was checked with a vacuum-tube voltmeter (Hewlett-Packard, Model 400-D). Both the 10 db and the 1 db settings gave accurate attenuation settings. The same test list recorded by a speaker was used at each of the five levels in succession, beginning with the most difficult level until all five listening conditions had been completed. For example, the listeners heard one speaker reading a randomized list of 80 vowels with the level of the auditory stimuli set 4 db above the average detection threshold of the listeners. Answer sheets were immediately collected, the tape was rewound, and the listeners were given a short break to eliminate fatigue. Next, the level was set 8 db above detection threshold. Again, the papers were collected, the tape rewound, and this time the level was set at 12 db. The was repeated at the 16-db and the 20-db levels for a total of five levels.

Conditions for testing in noise were similar to those used for testing in quiet except that the tests were conducted in the Psycholinguistics Laboratory instead of the I.A.C. room. The ambient noise level in the laboratory during the testing was 42 db, this reading being taken from a sound-level meter (H.H. Scott, Model 410-B, C scale). It was assumed that the masking noise fed into the earphones was sufficient to mask the ambient noise in the room. When questioned, the subjects reported that the only sound they heard was that coming to them through the earphones. The administration of the test in noise was in other respects the same as the administration of the test in quiet, with the most difficult level being presented first and the next four levels becoming progressively easier.

In that "the same test was used at each of five levels in succession," the question could be raised as to how much learning at one level was carried to the next. Two major factors reduce the likelihood that the list was learned sufficiently to affect significantly the intelligibility values at succeeding levels. First, each list contained 80 items, each item isolated and free of context or meaning. Second, each list was presented at the most difficult level initially and at successively easier levels. The listener would immediately discover at each successive level that the stimuli are clearer. In such a circumstance, it was reasoned that he would find it psychologically more economical to rely on listening to clearer stimuli rather than upon memory of those stimuli he had heard under less favorable conditions while remembering as well their positions in a series of 80 unrelated items.

IV ANALYSIS OF DATA AND RESULTS

The responses of the listeners were tabulated and analyzed in a variety of ways in order to yield answers to questions pertinent to the purposes of the study. This section of the report is organized in terms of these questions and the results relevant to each question.

1. Do the vowels and diphthongs differ significantly in intelligibility?

Listener responses to the vowel and diphthong stimuli were tabulated separately for each speaker in each condition. The data on the male and female speakers, respectively, were combined and are presented in Table I. Each stimulus (vowel or diphthong) elicited a total of 2400 responses; eight speakers presented each vowel or diphthong five times under ten conditions to six trained listeners. The number and percentage of correct identifications of each stimulus were tabulated and are presented in Table II.

TABLE I -- CORRECT RESPONSES AT DIFFERENT LEVELS OF QUIET (Q) AND NOISE (N)

	LEVEL I		LEVEL II		LEVEL III		LEVEL IV		LEVEL V		TOTAL	TOTAL	TOTAL
	Q	N	Q	N	Q	N	Q	N	Q	N	Q	N	Q & N
MALE SPEAKERS	ɑ 58	ɪ 66	ɑ 85	ɔ 82	ɑ 108	ɑʊ 111	ɑɪ 118	ɑʊ 119	ɑʊ 119	ɑʊ 119	ɑ 480	ɑ 466	ɑ 946
	ɔ 45	ɛ 57	ɔ 82	ɑ 80	ɑʊ 107	ɑ 109	ɑʊ 117	ɛ 113	ɪ 117	ɑ 117	ɑʊ 463	ɑʊ 460	ɑʊ 923
	ɑʊ 45	ɑ 52	ɑɪ 78	ɑʊ 79	ɔ 99	ɔ 105	ɑ 113	ɔ 111	ɑɪ 116	e 116	ɑɪ 448	ɪ 454	ɔ 895
	ɪ 43	e 52	ɑʊ 75	ɪ 78	ɑɪ 97	ɪ 103	ɔɪ 112	ɑɪ 110	ɑ 116	ɔ 115	ɔ 444	ɔ 451	ɪ 858
	ɛ 40	i 39	ɛ 66	e 72	ɛ 92	ɑɪ 100	ɔ 108	ʌ 109	o 115	ɑɪ 115	ɛ 406	ɛ 446	ɑɪ 856
	ɑɪ 39	ɔ 38	ɪ 59	ɛ 70	ɪ 91	ɛ 96	e 106	ɑ 108	e 114	ɪ 112	ɪ 404	e 434	ɛ 852
	o 37	ʌ 38	e 49	ɑɪ 66	ɔɪ 79	e 96	ɪ 105	ɔɪ 100	ɝ 110	ɔɪ 112	e 378	ɑɪ 408	e 812
	æ 33	o 37	ʌ 49	ʌ 59	ʌ 79	ɔɪ 86	ɛ 102	e 98	ɔ 110	ɝ 111	o 368	ʌ 398	o 749
	e 32	ɑʊ 32	æ 46	o 56	e 77	ʌ 83	o 99	o 96	ʌ 108	ɛ 110	ʌ 351	o 381	ʌ 749
	ʌ 24	æ 22	ɝ 45	i 52	ɝ 76	o 82	ɝ 98	ɪ 95	ɛ 106	o 110	ɔɪ 349	ɔɪ 354	ɔɪ 703
	i 21	ʊ 18	o 42	ɔɪ 44	o 75	i 75	æ 92	ɝ 94	ɪ 106	ʌ 109	ɝ 346	i 343	ɝ 656
	ɝ 17	u 17	u 40	u 41	æ 66	ɝ 71	ʌ 91	i 91	æ 104	æ 101	æ 341	ɝ 310	æ 641
	u 16	ɑɪ 17	ɔɪ 27	æ 40	u 52	æ 58	u 85	u 83	u 97	u 95	u 290	æ 300	i 585
	ʊ 16	ɪ 12	i 26	ɝ 29	i 51	u 54	i 68	æ 79	u 76	ɪu 94	i 242	u 290	u 580
	ɪ 14	ɪu 8	ʊ 21	u 26	ʊ 37	ɪu 37	ʊ 62	ɪu 66	i 76	i 86	ʊ 212	u 222	u 434
	ɪu 8	ɝ 5	ɪu 7	ɪu 15	ɪu 12	ʊ 35	ɪu 31	ʊ 60	ɪu 61	u 83	ɪu 119	ɪu 220	ɪu 339
Total	488	510	797	889	1198	1301	1507	1532	1651	1705	5641	5937	11578
FEMALE SPEAKERS	e 50	e 54	ɛ 85	ɪ 88	ɛ 100	ɛ 101	e 110	ɛ 107	ɔɪ 114	ɛ 113	ɛ 447	ɛ 444	ɛ 891
	ɪ 49	ɪ 47	e 76	e 82	e 96	ɪ 95	ɛ 108	ɪ 104	e 113	ʌ 111	e 445	ɪ 442	e 874
	ɑ 47	ɛ 45	ɪ 70	ɛ 78	ɪ 89	e 89	ʌ 105	ʌ 104	ʌ 112	ɪ 108	ʌ 420	e 429	ɪ 860
	ʌ 45	i 40	ʌ 70	ɑ 64	ʌ 88	ʌ 85	ɪ 101	e 97	ɛ 111	e 107	ɪ 418	ʌ 387	ʌ 807
	ɛ 43	ɑ 39	ɑɪ 69	ʌ 56	ɑ 86	ɑ 81	ɑɪ 100	ɑ 93	ɑɪ 110	ɑʊ 106	ɑɪ 407	ɑ 381	ɑ 770
	ɑɪ 42	ʌ 31	ɑ 68	i 55	ɑɪ 86	i 77	ɔɪ 98	ɪ 87	o 110	ɔɪ 106	ɑ 389	i 351	ɑɪ 724
	i 40	æ 29	o 49	ɑɪ 53	o 83	ɔɪ 68	o 91	o 87	ɪ 109	ɑ 104	o 354	ɑɪ 317	o 648
	æ 32	ɑɪ 17	æ 48	æ 44	æ 73	ɑɪ 62	ɑ 90	i 86	ɝ 107	ɑɪ 104	æ 337	o 294	i 647
	o 21	ʊ 14	i 45	o 42	i 65	o 61	æ 87	ɑʊ 85	ɑʊ 103	o 94	ɔɪ 318	ɔɪ 286	ɔɪ 604
	ɔ 15	ɑʊ 10	ɔ 37	ɔ 27	ɔ 64	ɑʊ 58	ɑʊ 80	ɔ 82	ɑ 98	i 93	i 296	ɑʊ 280	æ 604
	ɑʊ 11	o 10	ɔɪ 37	ɑʊ 21	ɔɪ 63	ɔ 55	ɔ 79	ɑɪ 81	æ 97	ɔ 87	ɔ 291	æ 267	ɑʊ 559
	ʊ 10	u 9	ɑʊ 25	ɔɪ 20	ɑʊ 60	æ 44	ɝ 76	u 67	ɔ 96	u 85	ɑʊ 279	ɔ 257	ɔ 548
	u 9	ɔ 6	u 24	u 18	u 54	u 41	u 73	æ 66	u 88	æ 84	u 248	u 219	u 467
	ɪu 8	ɔɪ 5	u 18	u 17	ɝ 39	ɪu 34	i 65	ɪu 63	i 81	ɝ 80	ɝ 239	ɪu 187	ɝ 397
	ɝ 6	ɝ 4	ɝ 11	ɪu 11	ʊ 19	u 30	u 35	ɝ 48	u 79	ɪu 76	u 161	ɝ 158	u 302
	ɔɪ 6	ɪu 3	ɪu 4	ɝ 10	ɪu 12	ɝ 16	ɪu 23	u 28	ɪu 33	u 51	ɪu 80	u 141	ɪu 267
Total	434	363	736	686	1077	997	1321	1285	1561	1509	5129	4840	9970
Total	922	873	1533	1575	2275	2298	2828	2817	3212	3214	10770	10777	21547

TABLE II

NUMBER AND PERCENTAGE OF CORRECT IDENTIFICATIONS FOR 2400 PRESENTATION OF EACH OF 16 VOWELS OR DIPHTHONGS

Vowel or Diphthong	Correct Identifications	
	Number	Per Cent
ɛ	1743	72.6
ɪ	1718	71.6
ɑ	1716	71.5
e	1686	70.3
ɑɪ	1580	65.8
ʌ	1556	64.8
ɑu	1482	61.8
ɔ	1443	60.1
o	1397	58.2
ɔɪ	1307	54.5
æ	1245	51.9
i	1232	51.3
ɝ	1053	43.9
ʊ	1047	43.6
u	736	30.7
ɪu	606	25.3

The hypothesis of no difference between the frequencies of correct identification was tested by means of a Chi Squared One Sample Test. This yielded a value of 1327 with 15 degrees of freedom; a value of 37.70 is needed for .001 significance. It can be concluded that the intelligibility of the vowels and diphthongs, measured by per cent correct identification, are very significantly different.

Additionally, the data were analyzed to determine which of any two vowels or diphthongs is significantly more intelligible under a variety of conditions. Previously described, these varied conditions involved noise or quiet, differing signal-to-noise ratios or signal intensities, and sex differences in speakers. There were 20 possible variations or combinations of these conditions; listeners heard a male or female speaker, in noise or quiet, at each of five levels. Thus, for each vowel or diphthong there were 20 intelligibility values obtained under these differing conditions.

For any given pair of stimuli (vowel or diphthong) there were 20 related pairs of intelligibility values. For example, [ɑ] and [ɔ] had 58 and 45 correct identifications when heard at level 1 in quiet from male speakers (see Table I). The set of 20 such related values was tabulated for each possible pair of stimuli; there were 120 such sets of related values. Wilcoxin's Test for Matched Pairs, for each pair of vowels or diphthongs, was computed. The resultant 120 _T_ values, along with their significances, are presented in Table III.

TABLE III

WILCOXIN _T_ VALUES FOR THE DIFFERENCES IN INTELLIGIBILITY OF PAIRS OF VOWELS OR DIPHTHONGS

	ɛ	ɪ	ɑ	e	ɑɪ	ʌ	ɑu	ɔ	o	ɔɪ	æ	i	ɝ	ʊ	u	ɪu
ɛ		84.5	92	59.5	58	9.5*	59	54	5**	15**	0**	0**	4**	0**	0**	0**
ɪ			102.5	80	67	28**	61	51*	10**	18**	0**	0**	4**	0**	0**	0**
ɑ				97.5	27.5*	55.5	45*	6.5**	8.5**	20**	0**	1**	3**	0**	0**	0**
e					72	28**	66.5	53.5	6**	20.5**	1**	0**	0**	0**	0**	0**
ɑɪ						99	95	49	22.5**	21.5**	6.5**	31.5**	0**	0**	1**	0**
ʌ							89.5	63.3	32**	24.5**	7**	9.5**	9.5**	0**	0**	0**
ɑu								67.5	69.5	48.5	54.5	59	1.5**	0**	3**	0**
ɔ									95.5	74.5	52.5	64	6**	1**	2**	0**
o										63	35.5**	52*	9.5**	0**	1**	0**
ɔɪ											84	88.5	20.5**	23.5**	15**	1.5**
æ												100	46*	11.5**	0**	0**
i													65	51*	0**	1**
ɝ														102	33*	28.5**
ʊ															10**	0**
u																54
ɪu																

*Significant at .05
**Significant at .01

In each row of the table, one can see which vowels or diphthongs are significantly less intelligibile than the vowel or diphthong which heads that row. The columns of the table reveal which vowels or diphthongs are significantly more intelligible than the vowel or diphthong which heads the column. For the vowel [o], for example, the vowels and diphthongs [æ], [i], [ɝ], [u], [u̇] and [ɪu] are significantly less intelligible, and the vowels [ɛ], [ɪ], [ɑ], [e], [aɪ], and [ʌ] are significantly more intelligible. It is interesting to note that 84 of the 120 pairs of stimuli are significantly different in intelligibility. Two conclusions seem warranted. As previously confirmed, the vowels and diphthongs, as a group of stimuli, are significantly different in intelligibility. Secondly, this significant difference is not the result of one or two quite unintelligible vowels among a majority of vowels about equally intelligible. Rather, there are significant differences between most pairs of vowels or diphthongs. This implies that there is a fairly stable order of intelligibility of the vowels and diphthongs.

2. Is the rank order of vowel and diphthong intelligibility similar in quiet and in noise?

The fact that vowels and diphthongs differ in intelligibility having been determined, data were analyzed to determine if the rank order of intelligibility was similar in noise and in quiet at each of the levels previously described. Essentially, do the correlations between noise and quiet rankings increase or decrease as the stimulus becomes clearer? The intelligibilities of the vowels and diphthongs in noise and quiet were ranked for each stimulus level and rho coefficients were computed. For levels 1 through 5 consecutively, the obtained values were: .90, .93, .93, .72, and .71. Two observations can be made from these findings. First, the rank orders of vowel-diphthong intelligibility in quiet and in noise are significantly related. Second, it appears that as the stimuli (vowel and diphthong) increase in clarity, the relationship between rank orders declines. The significances of the differences between these correlation coefficients do not reach the .05 level but do approach it.

3. Does the relative intelligibility of the vowels and diphthongs remain the same as the intensity of the stimuli are successively increased?

The stability of the rank order of vowel-diphthong intelligibility was investigated between the five levels involving successive increases in signal-to-noise ratios or signal intensity. As the clarity of the stimulus increases, the vowels and diphthongs, of course, become more intelligible. However, it is of interest to know if the rate of increase in intelligibility is similar. Do some vowels or diphthongs increase in intelligibility at a more rapid rate than others? The means of answering this question was to compute correlation coefficients between the various levels of stimulus presentation. If the rate of increase in intelligibility was similar for these vowels and diphthongs, the correlations between the levels should be uniformly high.

The data in quiet and in noise were combined; previous results revealed that the rank orders in quiet and in noise were quite similar. The numbers of correct identifications of the vowels and diphthongs for each level were correlated with every other level. Rho coefficients were obtained and are presented in Table IV.

Several observations can be made from an examination of Table IV. First, the intelligibilties of the vowels and diphthongs do not improve at a uniform rate as the listening conditions become easier. Some stimuli, for example the diphthong [ɔɪ] (see Table I), are found to have substantially improved relative intelligibility as listening becomes easier, while others such as [ɔ] bcome relatively less intelligible. (Decreases in the relative intelligibility of certain stimuli must accompany increases for other items; however, this does not imply that the percentage of correct identifications of any stimulus actually decreases under improved listening conditions.) Second, the greatest relationship between orders of vowel-diphthong intelligibility is found among adjacent levels of stimulus clarity. The further removed the levels, the less the relationship. These two observations suggest the conclusion that, in quiet, noise, or quiet and noise combined, the order of vowel-diphthong intelligibility at one level of listening difficulty would not be a good predictor of the order at a very different level of listening difficulty.

TABLE IV

CORRELATIONS OF VOWEL AND DIPHTHONG INTELLIGIBILITY VALUES BETWEEN FIVE LEVELS OF STIMULUS CLARITY

LEVELS	1	2	3	4	5
1		.86	.79	.67	.36
2			.98	.90	.64
3				.92	.74
4					.86
5					

4. Is the rank order of vowel-diphthong intelligibility consistent among speakers?

The data were analyzed with respect to speaker effect on the relative intelligibility of vowels and diphthongs. Are there differences in the rank order of vowel-diphthong intelligibility as spoken by different speakers? Eight speakers were used to present the stimuli; the rank order of intelligibility was determined for each speaker in quiet and in noise. An average intercorrelation of ranks was computed for the eight rankings in quiet; a similar computation was made for rankings in noise. The obtained average correlations were .49 and .46 respectively, which are significant at the .01 level. These values expose significant relationship among different speakers with respect to the order of vowel-diphthong intelligibility.

5. Do male and female speakers produce similar intelligibility values for the vowels and diphthongs?

The speakers included four males and four females. This permitted a study of the effects of sex on vowel-diphthong intelligibility. Coordinate with the previous procedure, the intelligibility values of the male and female speakers were ranked and analyzed separately. The average intercorrelations of the rankings of the male speakers were .62 in quiet and .42 in noise. For female speakers, the corresponding values were .64 and .61. These generally larger values for males and females separately, when compared with the combined values of .49 and .46, shown above, suggest that speakers of the same sex produce the vowels and diphthongs more alike with

with respect to intelligibility than speakers of different sexes.

Additional differences were found between male and female speakers. Male speakers appear to be more intelligible in vowel and diphthong production, and particularly more intelligible in noise. The intelligibility values of the male and female speakers were summed separately for each level of stimulus presentation; this was done for quiet; for noise, and for these conditions combined. For each of these conditions, the male and female speakers provided two values for each of the 16 vowels and diphthongs. The Wilcoxin Test for Matched Pairs was computed for each of these sets of paired values. The results are presented in Table V.

TABLE V

RESULTS OF WILCOXIN TEST OF THE DIFFERENCE IN INTELLIGIBILITY VALUES BETWEEN MALE AND FEMALE SPEAKERS UNDER VARIOUS CONDITIONS

LEVEL	CONDITION QUIET	NOISE	QUIET AND NOISE
Level 1	$\bar{T}$ = 45.5	$\bar{T}$ = 12.5**	$\bar{z}$ = 2.47*
Level 2	$\bar{T}$ = 53.3	$\bar{T}$ = 25.5*	$\bar{z}$ = 2.03*
Level 3	$\bar{T}$ = 38.5	$\bar{T}$ = 7.5**	$\bar{z}$ = 3.00**
Level 4	$\bar{T}$ = 17 **	$\bar{T}$ = 6.5**	$\bar{z}$ = 4.15**
Level 5	$\bar{T}$ = 31.5	$\bar{T}$ = 8 **	$\bar{z}$ = 3.69**
All Levels	$\bar{T}$ = 42.0	$\bar{T}$ = 3 **	$\bar{z}$ = 3.44**

*Significant at .05 level
**Significant at .01 level

In all the conditions of presentation, the male speakers had, in varying amounts, superior intelligibility. However, it is only under conditions of noise accompanying the stimulus that the male speakers are consistently and significantly more intelligible. These findings do not appear to be due to differences in quality or to differences in loudness between male and female speakers. The methodology of the study included (1) an attempt to equate the male and female speakers for quality, and (2) an actual equating of all voices for loudness. Therefore, it appears that males more intelligibly communicate vowels and diphthongs than do females. The relative superiority of the male speakers in noise is even more strongly supported by the data.

Finally, it was noted (see Table I) that as a group male speakers are more intelligible in noise than in quiet, and female speakers are more intelligible in quiet than in noise. Though the conditions of quiet and noise were equated for difficulty for all speakers combined at each of five levels, male speakers are consistently superior in noise in comparison with quiet; female speakers are consistently superior in quiet in comparison with noise. These superiorities are significant at the .01 confidence level. Given a choice between speech transmissions through noise or an equally difficult condition of low signal intensity, the male voice probably will be more intelligible with the former and the female voice will be more intelligible with the latter condition. This difference of conditions which favor the male and female voice may be restricted to low-frequency stimuli such as vowels.

6. What are the principal confusions among the vowels and diphthongs under conditions of quiet and noise?

The substitutions of other vowels or diphthongs for the stimulus presented were tabulated for each of the 16 vowels and diphthongs. The tabulations for the conditions of quiet and noise were made separately. These are presented in Table VI.

The determination of principal confusions was handled by first hypothesizing that substitutions of one vowel for another could reasonably be explained as guessing behavior. If this null hypothesis can be rejected, then other explanations would need to be postulated. The most likely of these explanations was that two given vowels truly are confused.

If nothing but guessing behavior is assumed, then each of 15 errors for a particular stimulus is equally probable. Thus, of the total number of errors made for a given stimulus, each type of error should occur 6.67 per cent of the time except for sampling fluctuations. For example, there were 258 errors made for the vowel [ɛ] presented in quiet (see Table VI). Each of the 15 types of errors should occur as a pure guess 17.2 times, or 6.67 per cent of 258, except for deviations due to sampling.

The standard error of this expected 6.67 per cent for each type of error was computed. From this value, the .01 limits of the percentages one might obtain in samples, sizes of 258 were found. The upper limit, at the .01 level, was 10.67 per cent. Expressed as a number of errors, this would be 27.5. Error frequencies for the vowel [ɛ] in quiet which exceed 27.5 cannot be explained adequately as guessing behavior. In this case, the vowels [ɪ], [ɔ], and [æ] show significant departures from expected frequencies; that is, they are significant confusions for [ɛ].

The asterisks in Table VI indicate that there is at least one significant departure from the expected frequency of error (or confusion) for every stimulus presented to listeners. A total of 99 such significant error frequencies were found, indicating that guessing and sampling are not adequate explanations of these high frequencies. This led to the conclusion that a number of vowels and diphthongs are truly confused with each other.

Finally, it needs to be noted that the design of this study and the analysis described above provide an identification of the principal confusions, though not necessarily all confusions. This occurs because each potential confusion is not independent of the others. For example, in the last row of Table VI, the vowels [i] and [u] are the principal confusions for [ɪu]. These two confusions account for 549 of a total of 702 errors. The remaining 153 errors are distributed over the other 13 possible substitution errors. Any of these 13 possible errors would need to occur more than 63 times to be classified as a confusion. Thus, the principal confusions tend to mask other potential confusions; if the principal confusions were removed, others might well be exposed.

7. Do related pairs of errors produce similar confusability values?

When errors for vowels and diphthongs are examined, it can be noted that each error is one of a pair. For example, the vowel [a] can be an error for [ɛ], and [ɛ] can be an error for [a].

Analysis of the data (Table VI) supports the conclusion that paired errors produce somewhat similar confusability values. This was studied by means of rank order correlations. For example, the [a] was an error for the vowel [ɛ] in 10 instances, and the vowel [ɛ] was an error for the

TABLE VI

CONFUSION MATRIX FOR VOWELS IN QUIET (Q) AND NOISE (N)

		ε	ɑ	ɪ	e	ɑɪ	ʌ	ɑu	ɔ	o	ɔɪ	i	æ	ɝ	ʊ	u	ɪu	Errors	Omit	Correct
ε	Q		2	42*	26	2	27	--	28*	19	7	8	48*	12	24	6	7	258	89	853
ε	N		8	21	26	6	47*	9	22	28	8	5	51*	14	33*	--	--	278	32	890
ɑ	Q	3		1	3	36*	67*	15	35*	9	1	6	75*	4	3	8	1	267	64	869
ɑ	N	9		3	4	61*	55*	27	25	3	2	12	103*	4	3	2	2	315	38	847
ɪ	Q	47*	5		21	4	7	--	3	10	1	56*	4	9	42*	41*	16	266	112	822
ɪ	N	37*	1		11	4	6	1	3	14	1	36*	--	9	107*	21	9	260	44	896
e	Q	25	1	63*		2	8	3	1	37*	8	51*	8	21	27	28	6	289	88	823
e	N	22	1	42*		4	4	1	1	121*	20	14	3	43*	16	13	1	306	31	863
ɑɪ	Q	7	85*	3	2		27	76*	10	10	4	6	27	7	4	7	1	276	69	855
ɑɪ	N	19	94*	16	17		33	130*	10	7	3	5	62*	1	6	2	1	406	69	725
ʌ	Q	79*	62*	5	7	20		6	47*	26	3	4	79*	6	15	4	--	363	66	771
ʌ	N	121*	55*	6	6	18		10	39*	13	4	8	88*	5	4	2	1	380	35	785
ɑu	Q	4	86*	1	4	183*	36		13	14	2	3	25	3	5	10	--	389	69	742
ɑu	N	26	74*	9	24	151*	51*		14	20	9	11	29	4	2	3	1	428	32	740
ɔ	Q	19	109*	6	2	13	63*	20		29	3	1	118*	3	7	2	--	395	70	735
ɔ	N	52*	67*	2	9	17	102*	34		15	10	5	149*	1	--	1	--	464	28	708
o	Q	26	5	26	200*	4	17	1	2		8	18	10	9	34	35	7	402	76	722
o	N	33	2	31	357*	6	2	3	1		26	9	1	9	11	4	2	497	28	675
ɔɪ	Q	62*	3	35	91*	7	26	8	27	75*		5	14	17	65*	6	2	443	90	667
ɔɪ	N	115*	4	25	180*	7	15	8	18	104*		8	6	14	31	--	--	535	25	640
i	Q	18	10	54*	17	4	16	4	12	12	4		9	12	7	95*	38*	312	350	538
i	N	17	11	36*	13	15	9	11	12	18	2		4	7	7	120*	71*	353	153	694
æ	Q	40	127*	8	7	23	48*	15	125*	17	9	6		4	6	15	2	452	70	678
æ	N	49	192*	13	10	33	118*	21	124*	10	2	5		2	4	1	--	584	49	567
ɝ	Q	22	8	127*	121*	6	6	3	1	32	1	64	5		46	36	14	492	123	585
ɝ	N	31	2	159*	251*	2	3	4	4	83*	8	49	8		69*	10	11	694	38	468
ʊ	Q	79*	4	307*	18	1	15	2	5	23	4	37	3	27		36	6	567	95	538
ʊ	N	86*	1	431*	20	3	7	5	3	18	6	36	3	24		14	6	663	28	509
u	Q	24	10	59	9	7	10	4	3	2	2	446*	10	3	13		42	644	183	373
u	N	4	6	52	18	5	4	4	1	7	2	587*	8	7	7		62	774	63	363
ɪu	Q	24	10	70	21	5	17	7	4	6	2	439*	7	13	16	148*		789	212	199
ɪu	N	15	10	51	14	7	4	7	2	11	3	437*	11	7	11	112*		702	91	407
Total	Q	479	527	807	549	317	390	164	316	321	59	1150	442	150	314	477	142	6604	1826	10770
Total	N	636	528	897	960	339	460	275	279	472	106	1227	526	151	311	305	167	7639	784	10777

*Significant confusion at .01 level

vowel [a] 12 times; similarly the vowel [ɪ] was an error for the vowel [ɛ] 63 times, and the vowel [ɛ] was an error for [ɪ] 84 times. Fifteen such pairs of error frequencies were associated with each vowel or diphthong. For each of the 16 vowels or diphthongs, a rank order correlation was computed. The coefficients yielded are presented in Table VII.

TABLE VII

RELATIONSHIPS BETWEEN RELATED CONFUSIONS FOR EACH OF 16 VOWELS OR DIPHTHONGS

Vowel or Diphthong	Rank Order Correlation
ɛ	.79
ɪ	.72
ɑ	.90
e	.73
aɪ	.55
ʌ	.67
au	.52
ɔ	.70
o	.25
ɔɪ	.70
æ	.69
i	.50
ɝ	.68
u	.74
u	.65
ɪu	.76

Additional information was gained on the nature of vowel confusions by plotting each vowel at the intersection of the first two formants using the data obtained by Peterson and Barney.[12] The diphthongs, identified by broken lines to represent movement from one vowel position to another, were drawn somewhat out of position to permit the construction of confusion vectors. The intelligibility, the prinicpal confusions (A, B, C,), and the totals were then entered, resulting in Figure 2 and Figure 3.

The confusions observed in both quiet and noise follow the predominantly horizontal pattern that would be predicted from formant analysis, and are strikingly similar to the results Pickett[15] found with vowels in flat and high-frequency noises and to the results of Miller's study[14] in which a low-pass filter was used to deliberately remove the higher frequencies. In both of the latter studies the vowels were presented in a consonantal context.

Miller states: "Most of these (confusion) lines run horizontally, which is what we would expect if the vowels were projected onto the ordinate as a result of removing all information about their position on the abscissa. However, there are minor deviations from this rule: confusions between had and hud ([æ] and [ʌ]); between hud and hawed ([ʌ] and [ɔ]); and between head and hawed ([ɛ] and [ɔ]) should have occurred but did not.." He concludes: "The simplest explanation of these discrepancies is that hid, hood, head, and hud contained short vowels, whereas heed, who'd, had, hod, and hawed contain long vowels." While there are several differences between the study by Miller and the present study, one difference

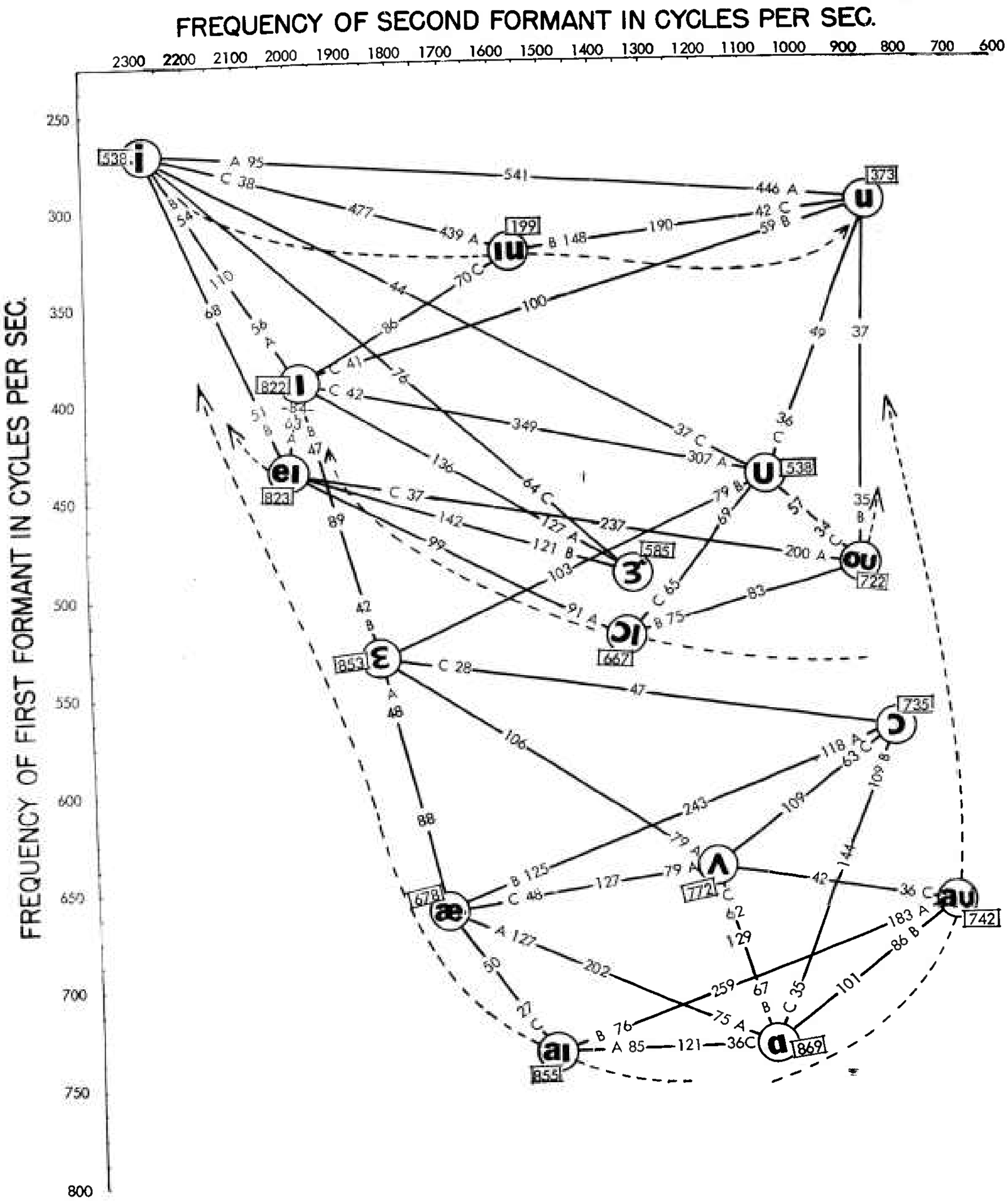


FIGURE 2

PRINCIPAL CONFUSIONS AMONG THE VOWELS AND DIPHTHONGS IN QUIET

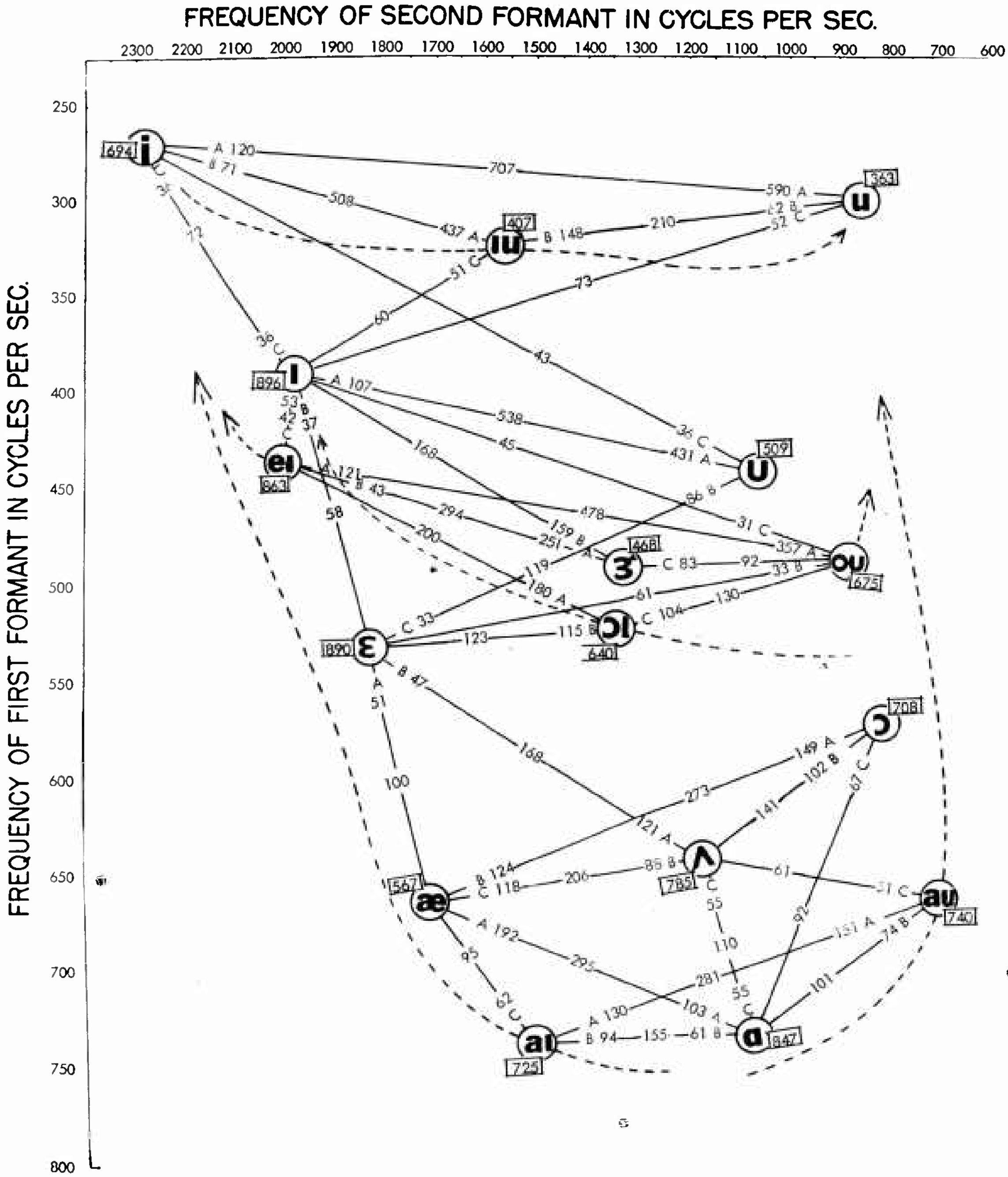


FIGURE 3

PRINCIPAL CONFUSIONS AMONG THE VOWELS AND DIPHTHONGS IN NOISE is that the vowels in the present study were sustained and therefore had similar temporal characteristics. The point of interest is that, in the present study, the confusions were found which Miller reported should have occurred, and these can be observed in Figure 2 and Figure 3.

V CONCLUSIONS

The data obtained relative to the principal questions raised in this study provide the bases for a series of conclusions; these are summarized below in the sequence of those questions listed in Section III.

1 a. The intelligibility values of the vowels and diphthongs are very significantly different at low intensities and relatively difficult signal-to-noise ratios.

b. The differences in vowel and diphthong intelligibility values reveal a fairly stable order of intelligibility. This order, from most to least intelligible, is [ɛ], [ɪ], [ɑ], [e], [aɪ], [ʌ], [aʊ], [ɔ], [o], [ɔɪ], [æ], [i], [ɝ], [u], [ʊ], and [ɪu].

2. The order of vowel and diphthong intelligibility under conditions of quiet and noise is highly related. The relationship is highest at the lowest levels of signal clarity.

3 a. As signal clarity improves, some vowels and diphthongs increase in intelligibility at a more rapid rate than others. Similarly, the converse is probably true; some vowels and diphthongs decrease in intelligibility more rapidly than others as the signal clarity declines. Therefore, the order of vowel-diphthong intelligibility at one level of listening difficulty would not be a good predictor of the order under a very different level of listening difficulty.

b. Conclusions (2) and (3a) suggest that changes in level of signal clarity affect the order of vowel and diphthong intelligibility more than change from quiet to noise or noise to quiet.

4. The rank orders of vowel-diphthong intelligibility among different speakers under similar conditions are significantly related, though not sufficiently for accurate prediction. Average correlations for quiet and noise conditions were .49 and .46.

5 a. Speakers of the same sex are more likely to have similar orders of vowel-diphthong intelligibility than speakers of different sexes.

b. Differences between male and female speakers in vowel-diphthong intelligibility was greater in noise than in quiet; the male speakers were superior in both conditions.

c. Female speakers were significantly more intelligible in quiet than in noise; male speakers were significantly more intelligible in noise than in quiet.

6 a. For each of the 16 vowels and diphthongs, at least one of the 15 other vowels or diphthongs is a significant confusion.

b. Approximately 40 per cent of all possible confusions, in this closed matrix of vowels and diphthongs, were found to be significant confusions. If these principal confusions are masking others, the number would still be greater.

c. The vowel [i] is the most frequently occurring confusion for another vowel or diphthong in both quiet and noise. Furthermore, the vowel [i] most frequently results in omission responses. It is interesting to note that in a study of the English digits,[17] the digit THREE, containing the vowel [i], is the least intelligible, and the most frequently confused with other digits.

7. Pairs of confusions are significantly related; for example, the relative frequency of [ɑ] as a substitute for [ɛ] is highly related to the relative frequency of [ɛ] as a substitute for [ɑ]. Confusions appear to bear a reciprocal relationship to one another.

BIBLIOGRAPHY

1. Fletcher, Harvey. Speech and Hearing in Communication. New York: D. Van Nostrand Company, 1953.

   Hirsh, Ira J. The Measurement of Hearing. New York: McGraw-Hill Book Company, 1952.

2. Harris, Cyril. "A Study of the Building Blocks in Speech." Journal of the Acoustical Society of America, XXV, September, 1953.

3. Sacia, C.F. and C.J. Beck. "The Power of Fundamental Speech Sounds," Bell System Technical Journal, V, 1926.

4. Black, John W. "The Effect of the Consonant on the Vowel," Journal of the Acoustical Society of America, X, January, 1939.

5. Fairbanks, G.G., A.S. House, and E.L. Stevens. "An Experimental Study of Vowel Intensities," Journal of the Acoustical Society of America, XXII, 1950.

6. Curry, E.T. "An Experimental Study of the Relative Identification Thresholds of Nine American Vowels," Speech Monographs, XVII, 1950.

7. Siegenthaler, B.M. "A Study of the Intelligibility of Sustained Vowels," Quarterly Journal of Speech, XXXVI, 1950.

8. Koch, W.E. and R.L. Miller. "Letters to the Editor: Dynamic Spectrograms of Speech," Journal of the Acoustical Society of America, XXIV, November, 1952.

9. Moser, H.M., J.J. Dreher, and J.J. O'Neill. The Masking of English Words by Prolonged Vowel Sounds, OSURF Project 664, Report 40, AFCRC TN 56-74, May, 1957.

10. Tiffany, William. "The Threshold Reliability of Recorded Sustained Vowels," Journal of Speech and Hearing Disorders, XVIII, December, 1953.

11. Lehiste, I., and Peterson, G.E. "Identification of Filtered Vowels," Phonetica, IV, 1959.

12. Peterson, G.E. and H.L. Barney, "Control Methods Used in a Study of the Vowels," Journal of the Acoustical Society of America, XXIV, March, 1952.

13. Edmondson, H.S. and E.M. Horwitz, "Cues for Vowel Discrimination," Journal of Speech and Hearing Disorders, XV, 1950.

14. Miller, George A. "The Perception of Speech," For Roman Jakobson, comp. Morris Halle et al, The Hague: Monton and Co., 1956.

15. Pickett, J.M. "Perception of Vowels Heard in Noises of Various Spectra," Journal of the Acoustical Society of America. XXIX, May, 1957.

16. Newman, Anthony K. A Laboratory-Standard Signal-to-Noise Equalizer, OSURF Project 1080, Report 59, AFCCDD TN 60-59, December, 1960.

17. Moser, H.M. and W.C. Fotheringham. Number Telling, OSURF Project 1080, Report 58, AFCCDD TN 60-40, December, 1960.

UNCLASSIFIED

UNCLASSIFIED